Characterisation and expression of the mitochondrial genome of a new type of cytoplasmic male-sterile sunflower

Spassova, Mariana; Moneger, Françoise; Leaver, Christopher J.; Petrov, Peter; Atanassov, Atanas; Nijkamp, H. John J.; Hille, Jacques

Published in:
Plant Molecular Biology

DOI:
10.1007/BF00019495

Publication date:
1994

Plant Molecular Biology **26**: 1819–1831, 1994.

Characterisation and expression of the mitochondrial genome of a new type of cytoplasmic male-sterile sunflower

Mariana Spassova[1,4], Françoise Moneger[2], Christopher J. Leaver[2], Peter Petrov[3], Atanas Atanassov[4], H. John J. Nijkamp[1] and Jacques Hille[1,*]

[1] *Department of Genetics, Institute for Molecular Biological Sciences, BioCentrum Amsterdam, Vrije Universiteit, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands (*author for correspondence);* [2] *Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK;* [3] *Institute of Wheat and Sunflower 'Dobroudja', 9520 General Toshevo, Bulgaria;* [4] *Institute of Genetic Engineering, Plant Biotechnology Research Centre, 2232 Kostinbrod-2, Bulgaria*

Received 8 March 1994; accepted in revised form 20 September 1994

Key words: cytoplasmic male sterility, *coxIII* and *atp6* loci, mitochondrial rearrangements, sunflower

Abstract

A new cytoplasmic male sterile sunflower, CMS3 [44], was characterised in relation to the *Petiolaris* (PET1) cytoplasmic male-sterile sunflower, CMS89 [25]. Southern blot analysis showed that the mitochondrial genome of CMS3 contains unique rearrangements in at least five loci (*atp6*, *atp9*, *atpA*, *nad1 + 5* and *coxIII*) compared to the PET1 sterile and the fertile cytoplasms. Transcripts of two (*coxIII* and *atp6*) of the five rearranged loci differed in CMS3 when compared to the corresponding loci in the PET1 and fertile cytoplasms. *In organello* protein synthesis experiments showed that the ca. 15 kDa mitochondrial polypeptide, characteristic of PET1, is not present in the CMS3 line. These data suggest that the molecular basis of male sterility in the CMS3 line differs from that of the PET1 cytoplasm. The nucleotide sequences of the coding and the immediate flanking regions of the *coxIII* and *atp6* genes of CMS3 were compared to the corresponding regions from the fertile sunflower. In CMS3, the ORFB-*cox* III locus is located immediately 3′ to the *atpA* gene whereas in the fertile cytoplasm these two loci are ca. 60 kb apart. This DNA rearrangement probably involved a 265 bp repeat which may be implicated in the DNA recombination associated with PET1 CMS. The *atp6* gene in CMS3 contains a 5′-terminal extention which results in an extended ORF. The potential involvement of the rearrangements associated with the *coxIII* and *atp6* loci in relation to the CMS phenotype is discussed.

Introduction

The mitochondrial (mt) genome of higher plants is viewed as a dynamic genetic system consisting of a heterogeneous population of molecules that interconvert via recombination involving pairs of direct and/or inverted repeats. The size and number of such repeats varies among species and

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers X82386 (*H. annuus* mitochondrial coxIII gene (ANTI)), X82387 (*H. annuus* mitochondrial atp6 gene (HA89)) and X82388 (*H. annuus* mitochondrial atp6 gene (ANTI)).

correlate with the complexity of the genome structure [27, 33]. Evidence for nonhomologous recombination has also been obtained in plant mtDNA with the finding of chloroplast DNA sequences integrated in the mitochondrial genome [45, 46].

Recombination plays an important role in introducing genetic diversity by generating mutants with altered mitochondrial genomes. One class of mitochondrial mutants is exemplified by cytoplasmic male sterility (CMS), a maternally inherited trait which affects pollen development. In the T-cytoplasm of maize intra- and intermolecular recombination events involving coding and noncoding regions have resulted in the formation of a novel chimeric gene (*urf13*), the protein product of which appears to be causally related to the CMS phenotype [8]. Recombination events involving the mitochondrial *at6*, *at9* and *coxII* genes have created chimeric sequences in the C male sterile cytoplasm of maize, which are also thought to be involved in the expression of the sterile phenotype [10]. The chimeric petunia gene *pcf* associated with CMS is also the result of mitochondrial genome rearrangements and contains part of the coding region of *atp9*, *coxII* and an unidentified reading frame termed *urf-S* [49].

In sunflower, CMS was first described by Leclercq [25] in the progeny of a cross between *Helianthus petiolaris* and *Helianthus annuus*. This type of male sterility known as PET1 CMS is the only one used commercially in hybrid sunflower seed production. In PET1 CMS the sterile phenotype is also associated with a rearrangement in the mitochondrial genome. The modified region of the mitochondrial genome is flanked by the *atpA* and *cob* genes and it has been suggested that a 265 bp repeat located at the 3′ end of the *atpA* gene and elsewhere in the genome may be involved in an inversion/insertion event leading to this genome rearrangement [22, 23, 41]. Due to this rearrangement downstream of the *atpA* gene, a new open reading frame ORF522 has been created which is co-transcribed with the *atpA* gene [22, 23]. A 15 kDa protein synthesised in mitochondria from the PET1 cytoplasm has been shown to be the translation product of ORF522 and may be responsible for the PET1 CMS phenotype [30].

It is important to increase the cytoplasmic genetic diversity in sunflower, by identifying and characterising new types of CMS, in order to reduce the dependence on a single cytoplasm and the associated risk of genetic vulnerability. Furthermore, characterisation of new types of CMS at the molecular level should allow a better understanding of the nuclear-mitochondrial interactions during plant development and particularly during microsporogenesis. We have initiated a molecular analysis of independently selected CMS lines, obtained either by specific crosses between different *Helianthus* species or by mutagenesis [44]. The molecular characteristics of three independent CMS lines appeared to be identical to PET1 CMS (Spassova *et al.*, in preparation). However, one of the genotypes investigated, obtained from an open pollination of *Helianthus annuus* ssp. *texanus* [48] and designated CMS3 showed a different mitochondrial genome organisation distinguishing this line from the parental fertile and the other CMS lines studied to date [44]. In the present article we report a detailed molecular analysis of the structure and expression of the mitochondrial genome of CMS3 in order to try and identify the genetic loci which cause male sterility.

Materials and methods

Plant material

Sunflower lines HA89 (fertile line) and CMS89 (CMS sterile line) are isonuclear lines containing the nuclear genotype of *Helianthus annuus*. The fertile line HA89 contains the cytoplasm from *H. annuus* and the sterile line CMS89 contains the cytoplasm from *H. petiolaris* and represents the PET1 CMS cytoplasm [26] which was used as a reference to study the new CMS3 line [44].

Table 1. Mitochondrial probes.

Code	Name	Origin	References
coxI	Cytochrome C oxidase subunit I	maize	[18]
coxII	Cytochrome C oxidase subunit II	maize	[13]
coxIII	Cytochrome C oxidase subunit III	*Oenothera*	[16]
cob	apocytochrome b	maize	[5]
rrn26	26S ribosomal RNA gene	wheat	[12]
rrn18	18S ribosomal RNA gene	wheat	[11]
rrn5	5S ribosomal RNA gene	wheat	[11]
atpA	subunit α of the ATPase complex	maize	[19]
atp6	subunit 6 of the ATPase complex	maize	[6]
atp9	subunit 9 of the ATPase complex	maize	[7]
nad1 + 5	NADH dehydrogenase subunit 1 and 5	*Arabidopsis*	[21]
nad3 + rps12	NADH dehydrogenase subunit 3 and part of the ribosomal protein 12	*Oenothera*	[15]
ORF873	*Sma* I-*Eco* RI 404 bp fragment located 3′ of the *atpA* gene in the fertile line	sunflower	[22]
ORF522	*Rsa* I-*Rsa* I 545 bp fragment located 3′ of the *atpA* gene in PET1 CMS	sunflower	[22]

Mitochondrial gene probes

The origins of the plant mtDNA sequences used as probes in hybridisation analysis are presented in Table 1.

Isolation of nucleic acids

Mitochondrial DNA and RNA were isolated from 10-day-old dark-grown sunflower seedlings. DNA was extracted as described by Köhler *et al.* [22] and RNA isolated according to the procedure of Stern and Newton [47], except that sucrose gradients were not used for the purification of mitochondria.

Southern and northern blot analysis

Mitochondrial DNAs digested with restriction endonucleases were separated by agarose gel electrophoresis and transferred to nylon membrane (Hybond N^+, Amersham) by vacuum blotting (LKB) according to the manufacturer's instructions. Mitochondrial RNA was fractionated by electrophoresis in a 1.3% (w/v) agarose gel containing 0.6 M formaldehyde and blotted to Hybond N^+ filters in the presence of 20× SSC. Mitochondrial DNA probes were labelled with [^{32}P]dCTP by random priming and hybridisations performed under the conditions described by Spassova *et al.* [44].

DNA sequence analysis

Cloning was carried out using vector pUC 119 and *Escherichia coli* strain JM 101. Ligation and transformation procedures were as outlined in Sambrook *et al.* [36]. DNA sequences were determined using a *Taq* Dye Primer Cycle Sequencing Kit (Applied Biosystems) and an automatic DNA sequenator (Applied Biosystems model 373A). Database searches were performed with FASTA [34].

In organello *protein synthesis*

Mitochondria were isolated and purified on sucrose gradient according to Leaver *et al.* [24] from 4-day-old dark-grown seedlings. *In organello* protein synthesis was achieved by incorporation

of [^{35}S]methionine into mitochondrial proteins following the method described by Leaver *et al.* [24]. Radiolabelled polypeptides were separated by electrophoresis on a 12–20% (w/v) SDS-acrylamide gradient gel which was dried and autoradiographed.

Results

Mitochondrial genome organisation

To identify regions of the mitochondrial genome in the CMS3 sunflower line that are potentially

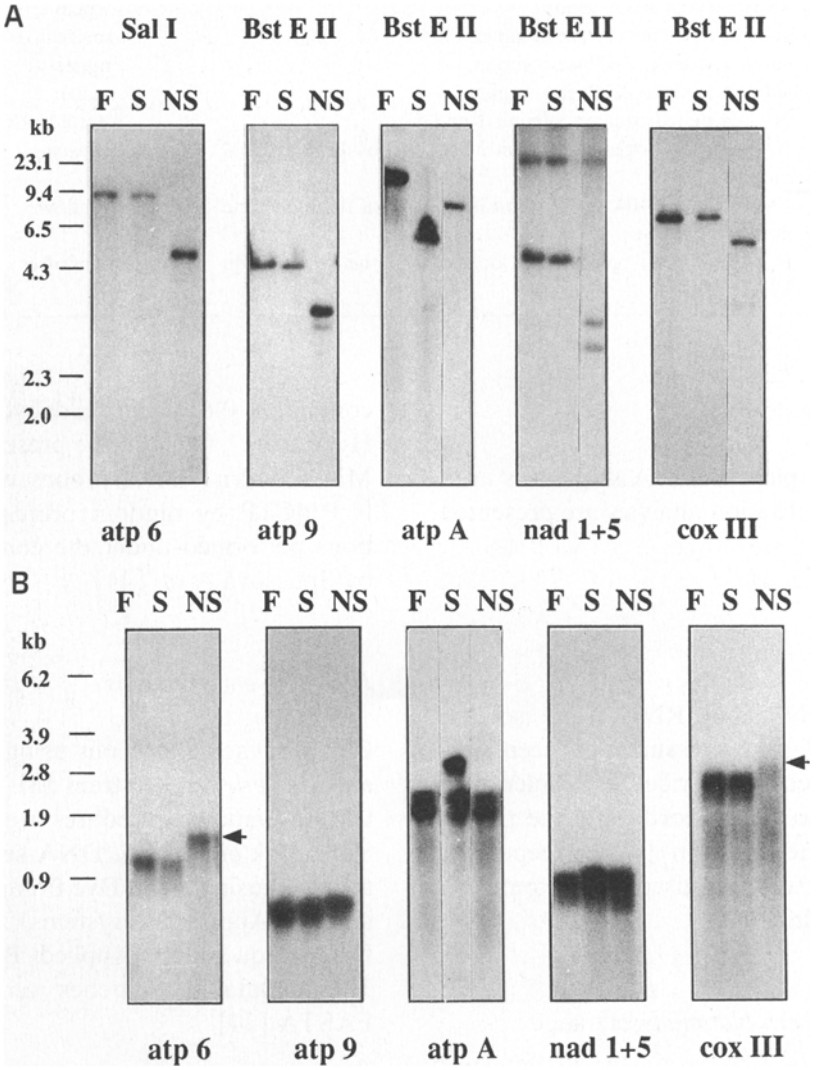


Fig. 1. A comparison of the organisation and expression of five mitochondrial gene loci in fertile and CMS lines of sunflower. A. Southern blot analysis: mtDNA isolated from young seedlings of fertile (F), PET1 CMS (S) and CMS3 (NS) sunflower lines was digested by restriction endonucleases, fractionated by electrophoresis and blotted onto nylon membrane and hybridised with gene-specific probes for the subunits 6 (*atp6*), 9 (*atp9*) and the α subunit (*atpA*) of the ATPase complex, the subunits 1 and 5 of the NADH dehydrogenase (*nad1 + 5*) and the subunit III of the cytochrome C oxidase (*coxIII*). Size markers are indicated on the left. B. Northern blot analysis: mtRNA isolated from young seedlings of fertile (F), PET1 CMS (S) and CMS3 (NS) sunflower lines, were fractionated by gel electrophoresis and hybridised to the same gene-specific probes as in Fig. 1A. Size markers are indicated on the left. Transcripts which differ in CMS3 compared to the fertile or PET1 CMS line are indicated by an arrow.

rearranged with respect to the fertile and PET1 CMS genomes, Southern blot analyses were carried out with a range of mitochondrial-specific gene probes. We have previously shown that an *atpA*-specific probe distinguished CMS3 from the fertile and PET1 CMS cytoplasms implying a rearrangement in the vicinity of the *atpA* gene is characteristic of CMS3 [41]. In order to further characterise genomic rearrangements associated with CMS3 sunflower, the investigation was expanded to include 14 additional mitochondrial genes widely distributed over the sunflower mitochondrial genome as hybridisation probes (Table 1). Only five of the gene probes *atp6*, *atp9*, *atpA*, *nad1 + 5* and *coxIII* showed differences in the restriction fragments to which they hybridised suggesting that the genomic environments of the five genes differ in CMS3 compared to the fertile and the PET1 CMS cytoplasm (Fig. 1A). Hybridisation with ORF873 and ORF522 probes [22] revealed that these ORFs located downstream of the *atpA* gene in fertile and PET1 CMS sunflower respectively are not present in the CMS3 mitochondrial genome. Interestingly the *atpA* and *coxIII* probes hybridised to the same *Hind* III and *Bgl* I fragments in the CMS3 line (data not shown), suggesting that due to recombination both genes are located close to each other in the CMS3 mitochondrial genome.

Comparisons of mitochondrial gene transcripts in mitochondria from fertile and sterile cytoplasms

The effect of mitochondrial genome rearrangements on transcription of individual genes was investigated by northern blot analysis of seedling RNA using the same mitochondrial gene probes as above. Transcript size and abundance of most of the genes investigated were similar in fertile and both CMS cytoplasms. The transcription pattern of the *atpA* gene in CMS3 is identical to the pattern in the fertile sunflower line where two major transcripts of ca. 2.3 and 2 kb were detected. The ca. 3 kb *atpA*-ORF522 chimaeric transcript specific to PET1 CMS is absent from the CMS3 and the fertile lines (Fig. 1B). Among the other rearranged loci in CMS3 only the transcripts of the *atp6* and *coxIII* genes differ from the fertile and PET1 CMS lines. The *atp6* probe hybridises to a 1.4 kb transcript in the fertile and PET1 CMS lines and to a larger transcript of 1.6 kb in the CMS3 line (Fig. 1B). The *coxIII* probe detects a transcript of 2.8 kb in the fertile and PET1 CMS lines and a diffuse transcript of about 3.1 kb in CMS3. When the same blot was reprobed with other mitochondrial genes discrete bands were obtained, indicating that the diffuse band is not the result of general RNA degradation.

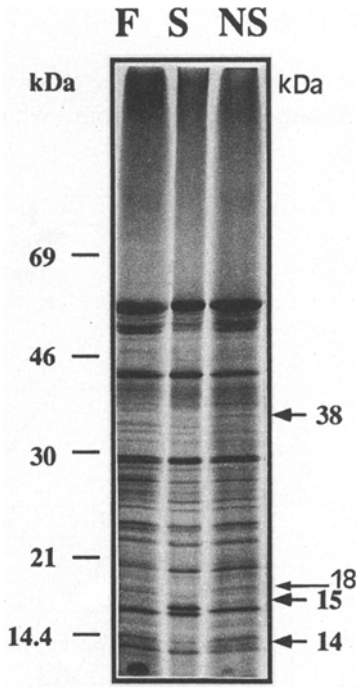


Fig. 2. A comparison of *in organello* mitochondrial protein synthesis by mitochondria from fertile and CMS lines of sunflower. Protein synthesis was performed using mitochondria isolated from dark-grown seedlings of fertile (F), PET1 CMS (S) and CMS3 (NS) sunflower lines. Mitochondrial proteins were fractionated by SDS-polyacrylamide gel electrophoresis and autoradiographed. Three proteins with a molecular mass of ca. 14, 18 and 38 kDa were only detected in the CMS3 line and are indicated by arrows. The 15 kDa protein specific to the PET1 CMS line is also indicated. Size markers are indicated on the left.

In organello *protein synthesis by mitochondria from sterile and fertile cytoplasms*

Analysis of the products of *in organello* protein synthesis by mitochondria isolated from seedlings of fertile, PET1 CMS and CMS3 lines revealed specific variations in mitochondrial translation products of the CMS3 cytoplasm. Mitochondria from CMS3 clearly do not synthesise the 15 kDa protein characteristic of the PET1 CMS. However, three other weakly labelled polypeptides with molecular masses of ca. 14, 18 and 38 kDa appeared to be specific to the CMS3 line and are not synthesised in mitochondria from either the fertile or the PET1 CMS lines (Fig. 2).

Cloning and sequencing of the coxIII *region of the mitochondrial genome*

In order to further characterise those regions of the CMS3 mitochondrial genome which are rearranged and show different expression than the fertile and PET1 CMS lines, the entire nucleotide sequence of the coding and immediate flanking regions of the *coxIII* and *atp6* genes in CMS3 was determined and compared with corresponding regions in the fertile line. A 7 kb *Hind* III restriction fragment from the CMS3 mitochondrial genome, carrying both the *atpA* and the *coxIII* gene was cloned and sequenced. The sequence data show that the recombination events which gave rise to the CMS3 cytoplasm brought together the *atpA* and *coxIII* loci (Fig. 3A).

Previous data have shown that a 265 bp repeat is present 3′ to the *atpA* gene and is thought to have mediated the DNA recombination events which led to the creation of the mitochondrial genome organisation characteristic of PET1 CMS (an 11 kb inversion and a 5 kb insertion event). This 265 bp repeat contains a portion of the 3′ region of the *atpA* open reading frame, its 3′-untranslated region and the beginning of the

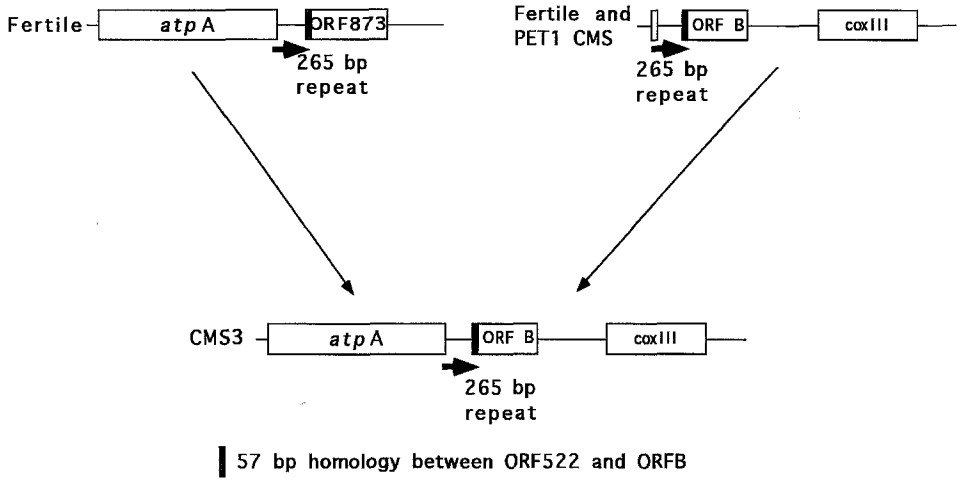
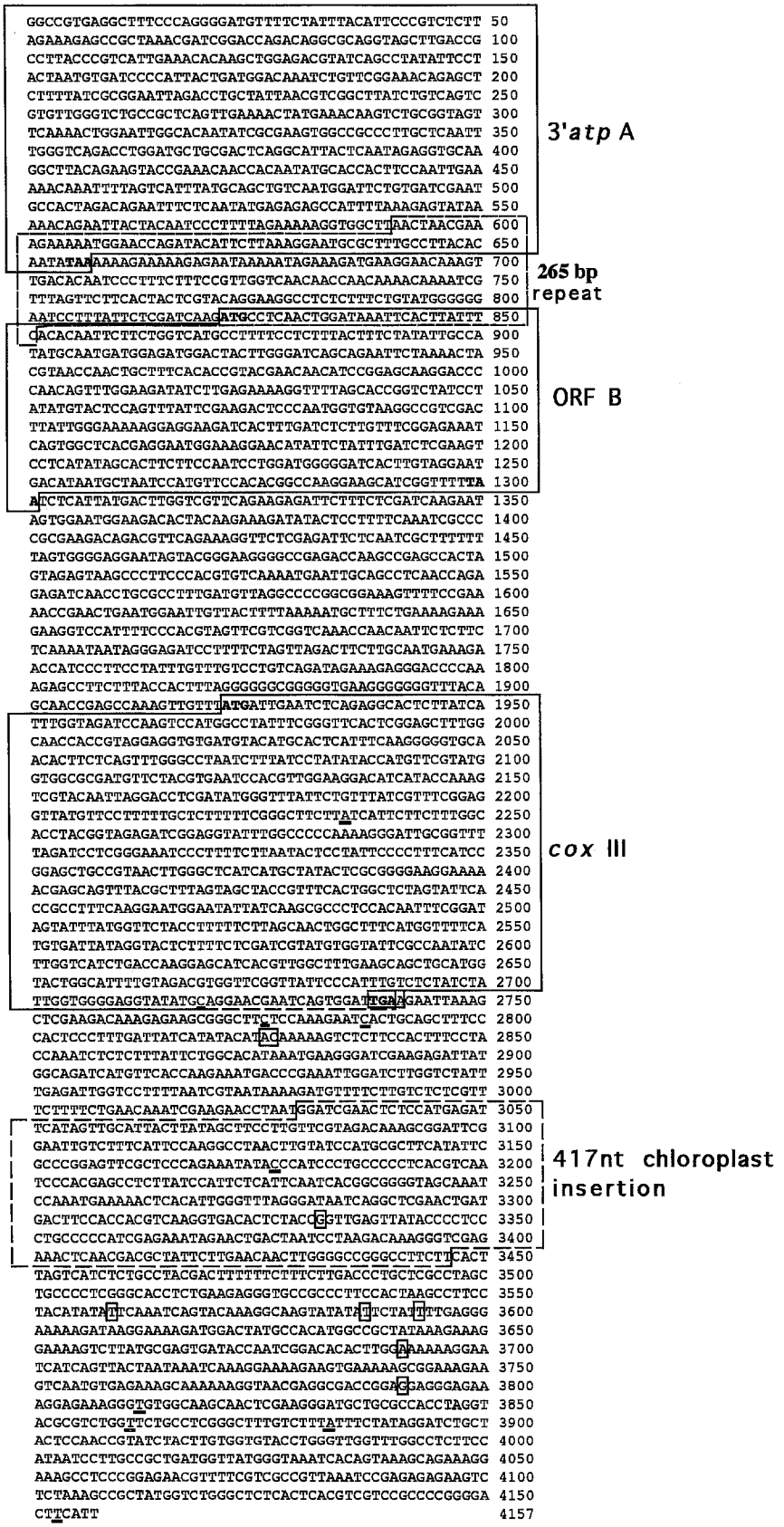


Fig. 3. Molecular organisation of the *coxIII* locus in CMS3 sunflower. A (this page). Schematic representation of the novel mitochondrial genome organisation of the *atpA-coxIII* region in CMS3 and its possible origins. As a result of rearrangement in the mitochondrial genome which led to the creation of CMS3, the *atpA* and *coxIII* loci are closely linked, in comparison to the fertile and PET1 CMS where they are at least 60 kb apart from each other [41]. Such a rearrangement probably involved the 265 bp repeats which are present elsewhere in the mitochondrial genome [23]. B (next page). Nucleotide sequence of the *atpA-coxIII* region in the CMS3 mitochondrial genome. The sequence shown covers the end of the *atpA* gene, the 265 bp repeat, ORFB, the *coxIII* gene and its 3′-untranslated region. Sequence comparison shows a complete conservation of the *atpA* gene and of the 265 bp repeat between CMS3 and the fertile genome. In the CMS3 line the DNA sequence continues with 99% identity to the ORF B-*coxIII* locus of the fertile sunflower. Open reading frames are boxed with a plain line and the 265 bp repeat is boxed with a dashed line. The nucleotide substitutions are underlined, the nucleotide insertions are boxed. The extension of the *coxIII* ORF by 7 amino acid residues is indicated by a dashed line. The 417 nt chloroplast DNA insertion located 3′ to the *coxIII* gene which is also present in the fertile mitochondrial genome is indicated.


3'atp A
265 bp repeat
ORF B
cox III
417nt chloroplast insertion

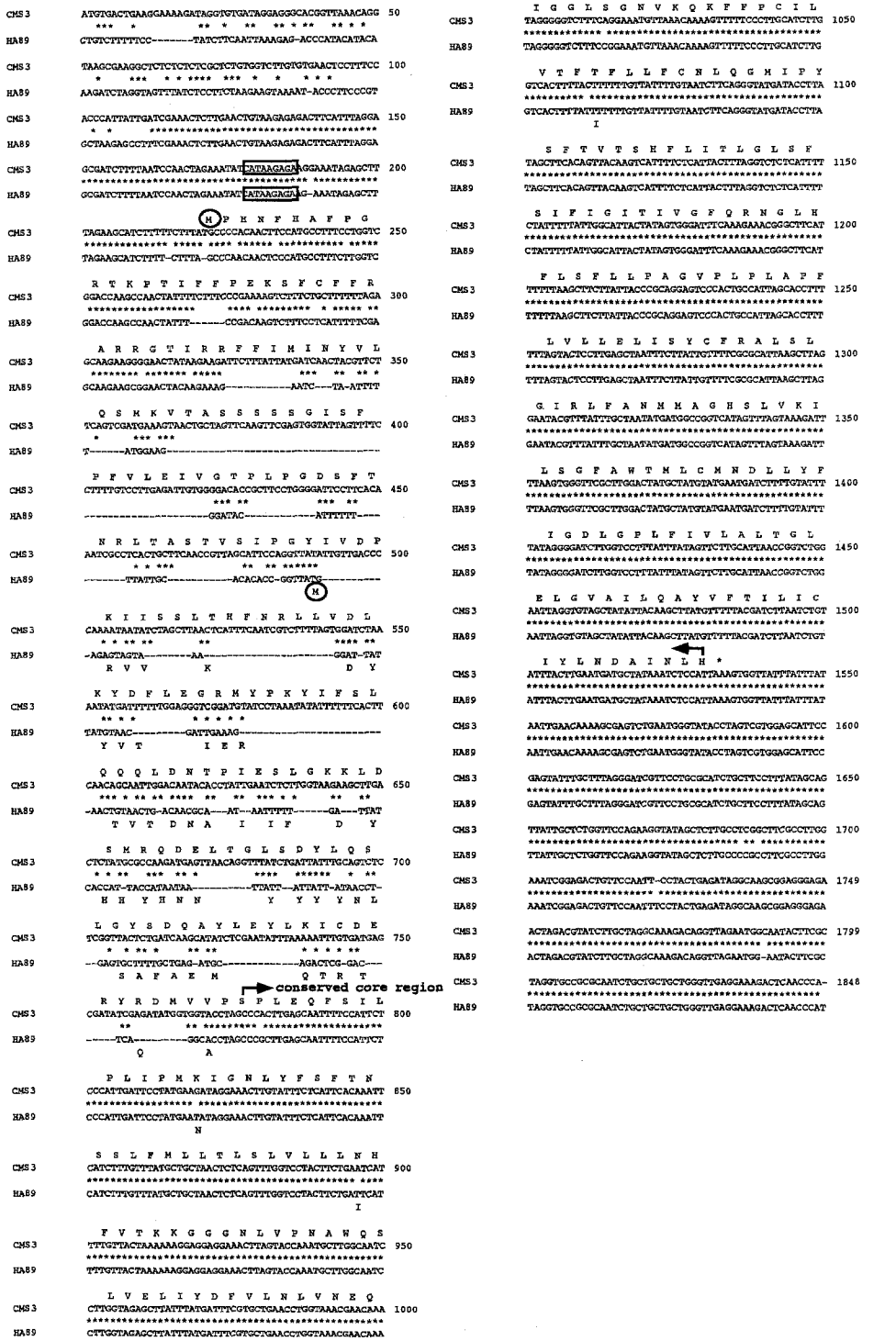


Fig. 4. A comparison of the molecular organisation of the *atp6* locus in CMS3 and in fertile sunflower. Alignment of the nucleotide and deduced amino acids sequences of the *atp6* genes from the CMS3 and fertile sunflower lines. Segments of the sequence which are 100% identical to the consensus putative plant mitochondrial promoter sequence as defined by Binder and Brennicke [2] are boxed. Asterisks represent conserved nucleotides between CMS3 and fertile HA89 line. Dotted lines represent gaps in the sequence alignment. Putative methionine initiation codons are circled. The conserved core region of 252 amino acids is indicated by arrows (see also Fig. 5).

ORF873 present in the fertile line. This 265 bp repeat is also present elsewhere in the fertile genome and includes the 5′ region of the ORFB which also shares the first 57 bp with ORF522 in the PET1 CMS genome [23]. ORFB, which has been shown to be cotranscribed with *coxIII* in the fertile line [35], has an unknown function but is conserved in other plant mitochondrial genomes. Interestingly, in the mitochondrial genome of CMS3, the *coxIII* locus is located immediately downstream of the *atpA* gene (Fig. 3) suggesting that the 265 bp repeat may have been involved in generating this novel rearrangement. According to Siculella and Palmer, the *atpA* and *coxIII* genes are at least 60 kb apart from each other in the fertile and PET1 CMS sunflower mitochondrial genomes [41]. Homology between the sequenced *atpA-coxIII* region in CMS3 and the corresponding regions of the fertile genome are almost 99%. The small number of sequence differences between the two lines were located immediately 3′ to the *coxIII* stop codon. These include single nucleotide substitutions and a few nucleotide insertions/deletions (Fig. 3B). Substitution of the nucleotide **A** at position 2719 (Fig. 3B) in the fertile sunflower with the nucleotide **C** in the CMS3 line results in elimination of the stop codon of the *coxIII* gene and leads to an extension of the predicted coding region of *coxIII* in CMS3 by seven amino acids. A 417 nt chloroplast DNA insertion is present in the mitochondrial genome of both the fertile and CMS3 lines 3′ to the *coxIII* gene [4].

ORFB in the mitochondrial genome of CMS3 is co-transcribed with the *coxIII* gene as in the fertile sunflower, but the ORFB-*coxIII* transcript in CMS3 is larger than in the fertile line (Fig. 1B). It is about 3.1 kb in size and exhibits a diffuse hybridisation pattern, indicating either transcript instability or multiple 5′ or 3′ ends.

Cloning and sequencing of the atp6 *region of the CMS3 mitochondrial genome*

The *atp6* loci of the fertile HA89 and CMS3 mtDNAs were cloned as 3.8 kb *Eco* RI and 3.0 kb *Bam* HI fragments respectively and the DNA sequences were determined. The sequences were aligned and are shown in Fig. 4. The *atp6* open reading frame in the fertile line, identified on the basis of nucleotide and amino acid sequence similarities with the *atp6* genes from other plant species, potentially encodes a protein of 301 amino acids (Fig. 5). The fragment containing the *atp6* region in the CMS3 line contains a 437 amino acids open reading frame which could encode a ca. 48 kDa polypeptide. Comparison of the *atp6* open reading frame in the fertile and the CMS3 line showed a conserved core region of 252 amino acids. This conserved core region is also found in the predicted ATP6 protein sequence from other plant species, as shown in Fig. 5. Upstream from the non-conserved/conserved sequence boundary (base 489, Fig. 4) the open reading frames of the *atp6* gene of fertile and CMS3 sunflower lines are extended by 49 and 185 codons respectively and

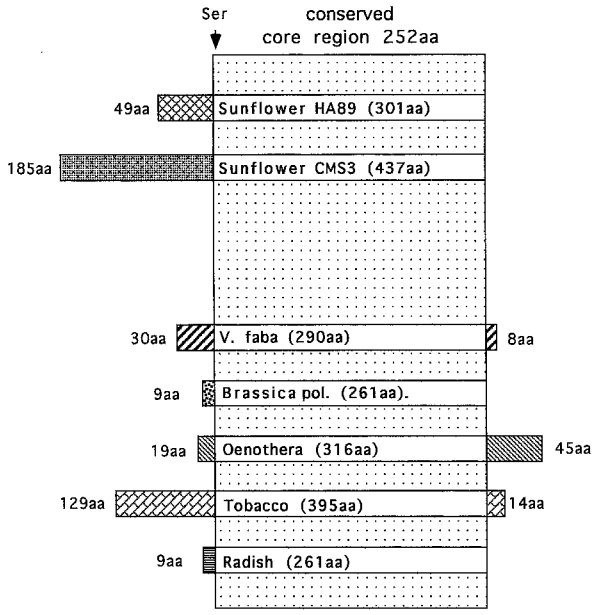


Fig. 5. Comparison of the predicted open reading frames encoding the ATP6 proteins in sunflower with different plant species. The open reading frame of the mitochondrial *atp6* genes vary extensively in size in different plant species. The arrow indicates the serine residue from which similarity between the different *atp6* sequences starts. A conserved core region of 252 amino acids can be identified. This is flanked by diverse amino and carboxy terminal extensions in different plant species.

show no similarity in the nucleotide or predicted amino acid sequences (Fig. 5).

Discussion

There is now strong evidence that CMS in maize (CMS-T cytoplasm), petunia and sunflower (PET1 CMS cytoplasm) is caused by the expression of single chimaeric genes, created as a result of aberrant recombination events [9, 30, 32]. However, in some cases such as CMS-C in maize [10] and CMS in carrot [37], rearrangements of several mtDNA regions have been associated with the CMS phenotype without direct proof of a causal relationship. In this paper, we report the molecular characterisation of a new type of CMS in sunflower. Southern blot analysis using 14 mitochondrial gene probes showed that the genomic environment around at least five mitochondrial loci has been altered in the newly identified CMS3 sunflower line (Fig. 1A). The number of rearrangements in the mitochondrial genome of the CMS3 line makes it difficult to establish a direct causal relationship between the gene(s) involved and the male sterile phenotype, since five loci were identified as potential candidates. An analysis of mitochondrial transcript patterns showed that among all five loci showing rearrangements in CMS3, only the transcripts of the *atp6* and *coxIII* loci differ from the fertile line (Fig. 1B). However, since our studies were carried out with RNA isolated from seedlings, changes in gene expression during anther development would not have been detected [43].

In organello protein synthesis by isolated mitochondria has shown that the novel 15 kDa polypeptide encoded by ORF522 in PET1 CMS line and associated with the CMS phenotype [17, 23], is absent in CMS3 (Fig. 2). This provides additional evidence that the molecular basis of male sterility in CMS3 is different from that in PET1 CMS. In order to further characterise the alterations in transcription and translation patterns in the CMS3 line, the unique mtDNA rearrangements associated with the *coxIII* and *atp6* loci were characterised.

The nucleotide sequence of the closely linked *atpA* and *coxIII* regions in CMS3 showed a high level of similarity with the corresponding, but widely separated, regions of the *atpA* gene and the *coxIII* gene in fertile and PET1 sterile sunflower (Fig. 3B) [41]. Since there is a 265 bp repeat at the 3′ end of the *atpA* gene and at the 5′ end of the ORFB-*coxIII* locus in the fertile sunflower [22] (Fig. 3A), we propose that a recombination event brought together the *atpA* and *coxIII* loci in the CMS3 line. If this is the case, another copy of the 265 bp repeat should be present 3′ of the *coxIII* gene, beyond the end of the sequence we have determined in this work but this remains to be verified. It is interesting to note that a probe specific for the ORF873, which is normally present 3′ of the *atpA* gene in the fertile line, does not hybridise to any restriction fragment in the CMS3 mtDNA. This suggest that after the DNA recombination events leading to the new mtDNA rearrangements found in CMS3 genome, the ORF873 is either absent or present at a very low level and therefore not detected in our southern blot analysis.

The novel rearrangement around the *coxIII* loci in CMS3 could have involved sequences responsible for transcript initiation of the ORFB-*coxIII* genes in CMS3 and thus affect the expression of these genes. In the CMS3 line we have shown that the *atpA* gene is not co-transcribed with the ORFB-*coxIII* genes as is observed in PET1 CMS with *atpA* and ORF522 [30], however, ORFB is co-transcribed with the *coxIII* gene in the CMS3 line, as is also the case in the fertile line (Fig. 1B) [35]. The fact that a larger *coxIII* transcript is detected in CMS3 compared to the fertile sunflower (Fig. 1B) probably results from an extention in the 3′ region of the locus (Fig. 3B). The diffuse pattern of this transcript, as compared to the corresponding transcript in fertile and PET1 CMS lines could be explained by either multiple transcription initiation and termination sites and/or by instability of the transcript. Further experiments are needed in order to distinguish between these different possibilities. In the CMS3 line, several nucleotide substitutions, insertions and deletions were detected 3′ to the *coxIII* gene

compared to the fertile sunflower line. These alterations in the nucleotide sequence do not change the potential to form stem and loop structures, but minor changes in the secondary structure of the *coxIII* locus could be functionally important as these might play a role in gene expression [39]. As in the fertile line, a 417 nt chloroplast DNA insertion is located 3′ to the *coxIII* gene (Fig. 3A). Whether this sequence plays a role in the expression of the *coxIII* gene is not known. Our characterisation of the genomic rearrangements and expression of the ORFB-*coxIII* locus in the CMS3 line compared to the fertile line suggest that these rearrangements could modify expression of the gene in the CMS3 line and be causally related to the CMS phenotype.

Sequence analysis of the *atp6* loci from both fertile and CMS3 lines revealed that the open reading frames encoding ATP6 varies extensively in size and could potentially encode a polypeptide of 301 amino acids in the fertile line and 437 amino acids in CMS3 (Figs. 4 and 5). The open reading frames of the *atp6* gene in both lines share a conserved core region of 252 amino acids and differ in the amino terminal region in that the ORF in the fertile line has an extension of 49 amino acids compared to an extension of 185 amino acids in the CMS3 line (Fig. 5). The amino terminal extension in both lines show little similarities to each other. Comparison of the predicted ATP6 amino acid sequences from *Vicia faba* [28], *Brassica polima* [42], *Oenothera* [40], tobacco [3] and radish [29] share the conserved core region of 252 amino acids (Fig. 5) but show diversity in the length of amino terminal and carboxy terminal extensions. The amino terminal extensions of higher plant ATP6 polypeptides predicted from DNA sequence data varies in length from 9 amino acids in radish to 129 amino acids in tobacco and the carboxy terminal extensions vary in length from zero in *Brassica polima* and radish to 45 amino acids in *Oenothera*. Although it has been suggested that the ATP6 polypeptides may contain presequences involved in membrane localisation [3] and proteolytic processing and maturation of the protein [31, 10], the role of these sequences as well as the identity of amino acids required for correct processing are unknown. In yeast mitochondria, the ATP6 protein can tolerate amino acid substitutions at many positions without a significant loss of function [20]. Therefore it is not clear whether the extended form of the ATP6 protein predicted from the DNA sequence data in CMS3 would be correctly localised and processed and whether the amino terminal extension has any effect on the function of the protein.

It should be noted that although there are several examples of chimeric genes and pseudogenes in plant mitochondria [1, 8, 14, 38], in most cases there is also an intact copy of the gene elsewhere in the genome. Interestingly, the CMS3 line contains only one copy of the *atp6* gene, therefore this chimeric locus must encode a functional ATP6 protein. Other examples of altered single copy mitochondrial genes correlated with a male sterile phenotype are *coxI* in the 9E cytoplasm of sorghum [1], *atp6* in *Ogura* radish [29] and *atp6*, *atp9* and *coxII* in CMS-C in maize [10]. Thus, we can speculate that a modified form of the ATP6 polypeptide in the CMS3 line may not impair the activity of the F1-ATP synthase complex during most stages of plant growth and development but fail to meet the demands associated with mitochondrial function during microsporogenesis.

A comparison of the proteins synthesised by mitochondria isolated from fertile, PET1 CMS and CMS3 lines showed that the CMS3 line does not synthesise the 15 kDa protein characteristic of the PET1 CMS line but does synthesise three weakly labelled polypeptides with molecular masses of ca. 14, 18 and 38 kDa not found in the other lines. The relationship of these polypeptides to the altered expression of the *atp6* and *coxIII* loci described above has not been determined at this stage. The availability of a fertile restored CMS3 line will be extremely useful in establishing a correlation between our molecular data and the expression of the CMS phenotype. Very recently, restorer genes for the CMS3 line have been found in the wild species *Helianthus annuus* ssp. *texanus* (Iuoras, personal communication) and this will allow further investigation concerning the molecular basis of CMS in this newly

characterised line. The identification and use of a novel CMS source in sunflower breeding, which is different from PET1 CMS, will increase the cytoplasmic genetic variability and reduce the potential risk of genetic vulnerability associated with the use of a single source of male-sterile cytoplasm.

Acknowledgements

We gratefully acknowledge Drs C.S. Levings III, A. Brennicke, W. Schuster, V. Knoop and R. Köhler for providing us with mitochondrial probes. This work was supported in part by a short-term EMBO fellowship to M. S.

References

1. Bailey-Serres J, Hanson DK, Fox TD, Leaver CJ: Mitochondrial genome rearrangement leads to extension and relocation of the cytochrome C oxidase subunit I gene in sorghum. Cell 47: 567–576 (1986).
2. Binder S, Brennicke A: Transcription initiation sites in mitochondria of *Oenothera berteriana*. J Biol Chem 268: 7849–7855 (1993).
3. Bland MM, Levings III CS, Matzinger DF: The ATPase subunit 6 gene of tobacco mitochondria contains an unusual sequence. Curr Genet 12: 475–481 (1987).
4. Ceci LR, Saiardi A, Siculella L, Quagliariello C: A tRNA ^{Val}(GAC) gene of chloroplast origin in sunflower mitochondria is not transcribed. Plant Mol Biol 23: 727–736 (1993).
5. Dawson AJ, Jones VP, Leaver CJ: The apocytochrome b gene in maize mitochondria does not contain introns and is preceded by a potential ribosome binding site. EMBO J 3: 2107–2113 (1984).
6. Dewey RE, Levings III CS, Timothy DH: Nucleotide sequence of ATPase subunit 6 gene of maize mitochondria. Plant Physiol 79: 914–919 (1985).
7. Dewey RE, Schuster AM, Levings III CS, Timothy DH: Nucleotide sequence of Fo-ATPase proteolipid (subunit 9) gene of maize mitochondria. Proc Natl Acad Sci USA 82:1015–1919 (1985).
8. Dewey RE, Levings III CS, Timothy DH: Novel recombination in the maize mitochondrial genome produce a unique transcriptional unit in the Texas male-sterile cytoplasm. Cell 44: 439–449 (1986).
9. Dewey RE, Timothy DH, Levings III CS: A mitochondrial protein associated with cytoplasmic male sterility in the T-cytoplasm of maize. Proc Natl Acad Sci USA 84: 5374–5378 (1987).
10. Dewey RE, Timothy DH, Levings III CS: Chimeric mitochondrial genes expressed in the C male-sterile cytoplasm of maize. Curr Genet 20: 475–482 (1991).
11. Falconet D, Lejeune B, Quetier F, Gray M: Evidence for homologous recombinations between repeated sequences containing 18S and 5S ribosomal RNA genes in wheat mitochondrial DNA. EMBO J 2: 297–302 (1984).
12. Falconet D, Delorme S, Lejeune B, Sevignac M, Delcher E, Bazetoux S, Quetier F: Wheat mitochondrial 26S ribosomal RNA gene has no intron and is present in multiple copies arising by recombination. Curr Genet 9: 169–174 (1985).
13. Fox TD, Leaver CJ: The *Zea mays* mitochondrial gene coding cytochrome oxidase subunit II has an intervening sequence and does not contain TGA codons. Cell 26: 315–323 (1981).
14. Grabau E, Havlik M, Gesteland R: Chimeric organisation of two genes for the soybean mitochondrial ATPase subunit 6. Curr Genet 13: 83–89 (1988).
15. Grohmann L, Brennicke A, Schuster W: The mitochondrial gene encoding ribosomal protein S12 has been translocated to the nuclear genome in *Oenothera*. Nucl Acids Res 20: 5641–5646 (1992).
16. Heisel R, Schobel W, Schuster W, Brennicke A: The cytochrome oxidase subunit I and III gene in *Oenothera* mitochondria are transcribed from identical promoter sequences. EMBO J 6: 29–34 (1987).
17. Horn R, Köhler RH, Zetsche K: A mitochondrial 16 kDa protein is associated with cytoplasmic male sterility in sunflower. Plant Mol Biol 17: 29–36 (1991).
18. Isaac PG, Jones VP, Leaver CJ: The maize cytochrome C oxidase subunit I gene: sequence, expression and rearrangement in cytoplasmic male sterile plants. EMBO J 4: 1617–1623 (1985).
19. Isaac PG, Brennicke A, Dunbar S, Leaver CJ: The mitochondrial genome of fertile maize (*Zea mays* L.) contains two copies of the gene encoding the α subunit of the F1-ATPase. Curr Genet 10: 321–328 (1985).
20. John VP, Willson TA, Lennane AW, Nagley P: Biogenesis of mitochondria: DNA sequence analysis of mit$^-$ mutations in the mitochondrial oli_2 gene coding for mitochondrial ATPase subunit 6 in *Saccharomyces cerevisiae*. Nucl Acids Res 14: 7437–7451 (1986).
21. Knoop V, Schuster W, Wissinger B, Brennicke A: *Trans*-splicing integrates an exon of 22 nucleotides into *nad5* mRNA in higher plant mitochondria. EMBO J 10: 3483–3493 (1991).
22. Köhler RH, Horn R, Lössl A, Zitsche K: Cytoplasmic male sterility in sunflower is correlated with the co-transcription of a new open reading frame with the *atpA* gene. Mol Gen Genet 227: 369–376 (1991).
23. Laver HK, Reynolds SJ, Moneger F, Leaver CJ: Mitochondrial genome organisation and expression associated with cytoplasmic male sterility in sunflower (*Helianthus annuus*). Plant J 1: 185–193 (1991).
24. Leaver CJ, Hack E, Forde BG: Protein synthesis by iso-

lated plant mitochondria. Meth Enzymol 97: 476–484 (1983).
25. Leclercq P: Une stérilité mâle cytoplasmique chez le tournesol. Ann Amél Plant 19: 99–106 (1969).
26. Leroy P, Bazetoux S, Quetier F, Delbut J, Berville A: A comparison between mitochondrial DNA of an isogenic male sterile (S) and male fertile (F) couple of sunflower. Curr Genet 9: 245–251 (1985).
27. Lonsdale DM, Hodge TP, Fauron CMR: The physical map and organisation of the mitochondrial genome from the fertile cytoplasm of maize. Nucl Acids Res 12: 9249–9261 (1984).
28. Macfarlane JL, Wahleithner JA, Wolstenholme DR: A broad bean mitochondrial *atp6* gene with an unusually simple non-conserved 5′ region. Curr Genet 18: 87–91 (1990).
29. Makaroff CA, Apel IJ, Palmer JD: The *atp6* coding region has been disrupted and a novel reading frame generated in the mitochondrial genome of cytoplasmic male sterile radish. J Biol Chem 264: 11706–11713 (1989).
30. Moneger F, Smart CJ, Leaver CJ: Nuclear restoration of cytoplasmic male sterility in sunflower is associated with the tissue-specific regulation of a novel mitochondrial gene. EMBO J 13: 8–17 (1994).
31. Mullen JA, Pring DR, Kempken F, Ferguson J, Chase CD: *Sorghum* mitochondrial *atp6*: divergent amino extensions to a conserved core polypeptide. Plant Mol Biol 20: 71–79 (1992).
32. Nivison HT, Hanson MR: Identification of a mitochondrial protein associated with cytoplasmic male sterility in *Petunia*. Plant Cell 1: 1121–1130 (1989).
33. Palmer JD, Shields CR: Tripartite structure of the *Brassica campestris* mitochondrial genome. Nature 307: 437–440 (1984).
34. Pearson WR, Lipman DJ: Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 85: 2444–2448 (1988).
35. Quagliariello C, Saiardi A, Gallerani R: The cytochrome oxidase subunit III gene in sunflower mitochondria is co-transcribed with an open reading frame conserved in higher plants. Curr Genet 18: 355–363 (1990).
36. Sambrook J, Fritsch EF, Maniatis EF: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).
37. Scheike R, Gerold E, Brennicke A, Mehring-Lemper M, Wricke G: Unique patterns of mitochondrial genes, transcripts and proteins in different male sterile cytoplasms of *Daucus carota*. Theor Appl Genet 83: 419–427 (1992).
38. Schuster W, Brennicke A: Pseudocopies of the ATPase α-subunit gene in *Oenothera* mitochondria are present on different circular molecules. Mol Gen Genet 204: 29–35 (1986).
39. Schuster W, Heisel R, Isaac PG, Leaver CJ, Brennicke A: Transcript termini of messenger RNAs in higher plant mitochondria. Nucl Acids Res 14: 5943–5954 (1986).
40. Schuster W, Brennicke A: Nucleotide sequence of the *Oenothera* ATPase subunit 6 gene. Nucl Acids Res 21: 9092 (1987).
41. Siculella L, Palmer JD: Physical and gene organisation of mitochondrial DNA in fertile and male sterile sunflower. CMS-associated variation in structure and transcription of the *atpA* gene. Nucl Acids Res 16: 3787–3799 (1988).
42. Singh M, Brown GG: Suppression of cytoplasmic male sterility by nuclear genes alters expression of the novel mitochondrial gene region. Plant Cell 3: 1349–1362 (1991).
43. Smart CJ, Monéger F, Leaver CJ: Cell-specific regulation of gene expression in mitochondria during anther development in sunflower. Plant Cell 6: 811–825 (1994).
44. Spassova M, Christov M, Bohorova N, Petrov P, Dudov K, Atanassov A, Nijkamp HJJ, Hille J: Molecular analysis of a new cytoplasmic male sterile genotype in sunflower. FEBS Lett 297: 159–163 (1992).
45. Stern DB, Lonsdale DM: Mitochondrial and chloroplast genomes of maize have a 12 kb DNA sequence in common. Nature 299: 698–702 (1982).
46. Stern DB, Palmer JD: Extensive and widespread homologies between mitochondrial DNA and chloroplast DNA in plants. Proc Natl Acad Sci USA 81: 1946–1950 (1984).
47. Stern DB, Newton KJ: Isolation of plant mitochondrial RNA. Meth Enzymol 118: 488–496 (1986).
48. Vranceanu AV, Iuoras M, Stoenescu FM: A contribution to the diversification of cytoplasmic male sterility in sunflower. HELIA 9: 21–25 (1986).
49. Young EG, Hanson MR: A fused mitochondrial gene associated with cytoplasmic male sterility is developmentally regulated. Cell 50: 41–49 (1987).